OPEN ACCESS

Estimated mortality of adult HIV-infected patients starting treatment with combination antiretroviral therapy

Constantin Theodore Yiannoutsos,[1] Leigh Francis Johnson,[2] Andrew Boulle,[2] Beverly Sue Musick,[1] Thomas Gsponer,[3] Eric Balestre,[4] Matthew Law,[5] Bryan E Shepherd,[6] Matthias Egger,[3] for the International Epidemiologic Databases to Evaluate AIDS (IeDEA) Collaboration

▶ Additional data are published online only. To view these files please visit the journal online (http://dx.doi.org/10.1136/sextrans-2012-050658).

[1]Department of Biostatistics, Indiana University School of Medicine, Indianapolis, Indiana, USA
[2]Centre for Infectious Disease Epidemiology and Research, University of Cape Town, Cape Town, South Africa
[3]Institute for Social and Preventive Medicine, University of Bern, Bern, Switzerland
[4]ISPED, Centre INSERM U897-Epidemiologie-Biostatistique, Universitè Bordeaux, Bordeaux, France
[5]The Kirby Institute, University of New South Wales, Sydney, Australia
[6]Department of Biostatistics, Vanderbilt University, Nashville, Tennessee, USA

Correspondence to
Professor Constantin Theodore Yiannoutsos, Department of Biostatistics, Indiana University School of Medicine, 410 West 10th Street, Suite 3000, Indianapolis, IN 46202, USA; cyiannou@iupui.edu

UNAIDS Report 2012 Guest Editors
Karen Stanecki
Peter D Ghys
Geoff P Garnett
Catherine Mercer

Accepted 29 September 2012

ABSTRACT

Objective To provide estimates of mortality among HIV-infected patients starting combination antiretroviral therapy.

Methods We report on the death rates from 122 925 adult HIV-infected patients aged 15 years or older from East, Southern and West Africa, Asia Pacific and Latin America. We use two methods to adjust for biases in mortality estimation resulting from loss from follow-up, based on double-sampling methods applied to patient outreach (Kenya) and linkage with vital registries (South Africa), and apply these to mortality estimates in the other three regions. Age, gender and CD4 count at the initiation of therapy were the factors considered as predictors of mortality at 6, 12, 24 and >24 months after the start of treatment.

Results Patient mortality was high during the first 6 months after therapy for all patient subgroups and exceeded 40 per 100 patient years among patients who started treatment at low CD4 count. This trend was seen regardless of region, demographic or disease-related risk factor. Mortality was under-reported by up to or exceeding 100% when comparing estimates obtained from passive monitoring of patient vital status.

Conclusions Despite advances in antiretroviral treatment coverage many patients start treatment at very low CD4 counts and experience significant mortality during the first 6 months after treatment initiation. Active patient tracing and linkage with vital registries are critical in adjusting estimates of mortality, particularly in low- and middle-income settings.

INTRODUCTION

Thirty years after the first description of the AIDS, cases of the disease have been observed in every part of the world. More than 33 million individuals are infected with the majority of them living in low- and middle-income countries.[1] However, there are also positive statistics. According to the same report, new infections have been reduced by 19% since 1999. In addition, up to 5.2 million people in low- and middle-income countries were receiving care in 2009, 1.2 million of whom started therapy for the first time in that year; a significant 30% increase compared with only a year before. Still, this only represents just over a third of all persons, 15 million by some estimates, who need therapy now.

Regardless of the enormity of this challenge, it is obvious that we are witnessing a historic undertaking with rapid scale-up of services providing antiretroviral therapy (ART) to millions of HIV-infected individuals, particularly in resource-limited parts of the world. By all accounts, the largest pharmacological intervention in human history has reduced rates of opportunistic infections, increased survival and enhanced the quality of life of the recipients of these services.[2 3] Nevertheless, it is clear that the impact of therapy is not the same across all groups of patients living with HIV/AIDS or all regions of the world.[2] For example, mortality is higher among patients with greater suppression of their immune system as indicated by lower CD4 lymphocyte levels at the start of therapy despite access to treatment[4] and response to therapy may not be as strong compared with patients with higher CD4 counts.[5 6] Gender and age play a role as well. In resource-constraint settings, men generally access treatment later than women[7] and may be less compliant to therapy once initiated.[8–10] Also, the disease may affect younger and older patients differently.[11 12] Thus, assessment of treatment efficacy, programme evaluation and policy making should, at a minimum, take these disease-specific and demographic factors into account.

One core component for assessing programme and treatment effectiveness is the estimation of patient mortality, especially among those patients who initiate ART. These estimates will in turn inform policy makers, modellers, epidemiologists and other stakeholders involved with antiretroviral programmes throughout the world. There are a number of advantages in considering mortality as the primary patient outcome. Mortality is the ultimate endpoint by which to evaluate the effectiveness of an intervention when, as is the case with HIV infection, the disease represents a grave risk to the survival of the patient. Another advantage is the unequivocal nature of the death as an outcome (as opposed to diagnosing a bacterial infection, for example). However, significant challenges remain. In all settings, but particularly in low- and middle-income countries, a substantial number of patients starting therapy are lost from observation. In one review covering a number of countries in sub-Saharan Africa, between a quarter and half of all patients who started therapy were

lost after 2 years from starting therapy.[13] Two reports from patient cohorts in the US and the UK also cited high levels of patient loss[14 15] suggesting that this phenomenon is not limited to resource-limited settings. Moreover, it appears that, at least in sub-Saharan Africa, the risks for mortality overlap to a large extent with the risks of loss from follow-up[16 17] and that the hardest to reach patients may also be those at the highest risk of mortality.[18] Thus, it comes as no surprise that programmes which undertake more intensive efforts to locate patients who have been lost have higher levels of mortality.[2] The immediate consequence of all these considerations is that mortality estimates will be negatively biased when derived exclusively from information on deaths obtained passively among patients who were maintained under observation until the end. The result is a potentially significant underestimation of the true death rates in reports which are based solely on passive death information.

Recently, mathematical models originally developed to address issues of informative dropout in studies of survival have been adapted to the context of ongoing epidemiological cohorts with similar non-random losses from follow-up.[16 17 19] These efforts are emanating from the work on *double sampling*[20] by Frangakis and Rubin[21] and inverse probability of treatment weighting schemes introduced by Robins and others.[22 23] They are based on the availability of vital status information on a random sample of patients among those who have been lost to follow-up (the second or 'double' sample). The methodology weighs those lost patients with available vital status information by the inverse of their number over all patients who were lost. In this manner, a pseudo dataset is constructed where the actively sought patients, including all deaths thus ascertained, represent themselves as well as all other patients who were lost but were not sought. Patients who remain on observation or have died (and their death has been recorded passively, without resorting to outreach methods) are weighted by one (ie, they only represent themselves). In this manner, an adjustment of the mortality estimate is made. It can result in dramatically higher mortality estimates because the hazard of mortality among patients lost to follow-up can be over 10 times higher compared with the hazard among patients who remain under observation.[17 19]

Information on the vital status of a sample of patients who are lost to follow-up can be obtained in one of two ways. First, a random sample of the individuals who were lost to follow-up can be traced employing methods such as telephone calls and home visits to establish whether the patient is still alive. In sub-Saharan Africa, there are a number of cohorts employing patient outreach programmes, which trace a significant proportion of lost patients.[24 25] A second way to determine patient vital status is through national vital registration systems. The latter approach is only feasible if two conditions are met: (1) a reasonably high proportion of deaths are recorded and (2) that health facilities capture the information required for the purpose of linking the patient to the vital registration system. In sub-Saharan Africa, South Africa is one of few countries in which the first condition is met[26] with around 90% of all adult deaths recorded.[27] The second condition is met only in a subset of South African ART cohorts. In those cohorts, the lost patients who can be successfully located in the death registries can then be treated in exactly the same way as successfully located patients who are physically traced after loss to follow-up. This means that the patients with identity numbers who are lost to follow-up are 'weighted up' to represent other lost-to-follow-up patients without identity numbers.[28]

Unfortunately, access to death registries is the exception rather than the rule in low- and middle-income countries. In addition, there are few sufficiently documented outreach programmes. This is because the goal of most outreach programmes is to locate as many patients as possible in order to reinstate them into care rather than to perform a public health evaluation or attempt to mathematically adjust mortality estimates in their patient population. In many cases, the public health value of documenting patient encounters in the process of outreach is not clear to even mature care and treatment programmes. To overcome this problem, a number of intermediate or hybrid methods have recently been proposed. For example, when outreach data are not available but there is information about the HR between the observed versus lost patients (possibly through outreach programmes at sentinel sites in the same country or region), the mortality of the latter subcohort can be imputed.[29] In other cases, one may rely on the observation that programmes with high rates of losses to follow-up also have lower HRs for mortality between patients who are lost and those who remain under observation. This has been observed in a number of instances throughout sub-Saharan Africa where rapid scale-up of care and treatment services result in often chaotic patient self-referral patterns from one clinic to another, which accounts for an increasing proportion of losses to follow-up.[30] Using a structured survey of a number of programmes throughout the region, Egger and colleagues constructed a nomogram where a correction factor is estimated according to the mortality ratio between patients who were lost or not and the level of loss to follow-up at a particular programme.[31] This correction factor is then used to produce adjusted estimates of 1-year mortality.

Death is thus an unequivocal measure of effectiveness of the care and treatment provided, particularly in the context of a fatal disease. Patient survival also affects a number of downstream economic and HIV-related factors, such as disease prevalence and incidence rates, the productivity of entire societies and so on. Thus, accurate estimation of mortality impacts numerous aspects of scientific enquiry as well as public health policy. This survey paper uses established mathematical modelling methodologies to provide estimates of mortality in HIV-infected patients starting ART by assembling data from a number of cohorts of patients receiving routine medical care and ART in a number of programmes around the world. Although we provide brief description of the methods, our main goal is to provide as geographically complete coverage of the world as possible and thus describe the burden of mortality in a number of world regions.

METHODS

Description of cohorts involved in the study

In this paper, we report on five regional cohorts with available patient-level data needed for estimation of mortality in adult HIV-infected patients starting ART: East, Southern and West Africa, Asia Pacific and Latin America.[32–36] All five are members of the worldwide International Epidemiologic Databases to Evaluate AIDS (IeDEA) Collaboration, a consortium created to aggregate epidemiological data from HIV regional cohorts around the world. The first of these five IeDEA cohorts is a network of 23 main care and treatment clinics and several satellite clinics comprising the partnership between the United States Agency for International Development and the Academic Model Providing Access to Healthcare (the USAID-AMPATH partnership), located in western Kenya. This cohort includes data obtained from a large number of patients

initially lost from follow-up who were subsequently successfully located through AMPATH's massive patient outreach programme.[16 17] The second cohort is comprised of nine care and treatment facilities in the Republic of South Africa. These cohorts are unique in our study as they have well-documented linkages with vital registries and information recovered from these linkages has been used to adjust estimates of mortality in these cohorts.[28] The third, fourth and fifth cohorts comprise care and treatment programmes in West Africa[36] and Asia Pacific, made up of the Therapeutics Research, Education, and AIDS Training in Asia (TREAT Asia) Study to Evaluate Resistance and the TREAT Asia HIV Observational Database,[34 37] and three Latin American cohorts participating in the Caribbean, Central and South American Network.[33]

Inclusion criteria

Data from all cohorts were obtained through routine clinical care of HIV-infected patients. Inclusion criteria were minimal, specifying simply that the patients were adult (aged >15 years) HIV-infected men and women who were ART naive at enrolment (except for limited previous exposure to antiretrovirals for prevention of mother-to-child HIV transmission) and who initiated first-line ART in any participating site within the cohort. Further detail on prevailing first-line regimens at each region is given in the Results section in the subsection related to each of the five cohorts comprising this study. While second-line treatment regimens are becoming increasingly available in many low- and middle-income settings, the prevalence of second-line therapy is still relatively low, particularly within the sub-Saharan African regions. Even though information on treatment switches to second-line therapy was available in a number of cohorts, we chose not to include this information in the mortality calculations both for simplicity of the modelling and because our main objective in this study was to assess the levels of mortality of a typical HIV-infected patient starting first-line ART in these settings.

Data collected

Data included demographic information such as gender and age. Age was grouped into four categories spanning 15–24, 25–34, 35–44 and ≥45 years. In addition to information on gender and age at initiation of ART, data included all landmark dates of each patient's treatment history, such as the date of ART initiation, the date of the last clinic visit, the date of death and the date of the most recent outreach encounter when applicable. A number of indicators were included in the data such as whether the patient died, whether the patient death was ascertained via passive means or through patient outreach and whether the patient had been lost to follow-up. Loss to follow-up was defined as having no clinical visits for more than 6 months. Gaps in care, where a patient discontinued attending the clinic for some time but subsequently returned to care, were not considered as losses to follow-up. On the other hand, patients who were outreached were essentially treated as having been lost to follow-up regardless of the length of time between their last visit and the outreach encounter date as, once traced successfully, their loss to follow-up status would never be determined with certainty.[17] This only applies to cohorts with documented outreach programmes (Kenya) or established linkages to vital registries (South Africa).

CD4 count at ART initiation was defined as the CD4 count available within 6 months before (strictly speaking 182 days before) and 14 days after initiation of therapy (the latter to ensure inclusion of CD4 counts which were obtained contemporaneously with the start of treatment but were recorded with a small delay). If there were multiple CD4 measurements during this period, then the one closest to the start date of ART was selected. The CD4 categories used were <50, 50–99, 100–199, 200–249, 250–349, 350–499 and ≥500 cells/μl.

Mathematical models of patient mortality

We adjusted mortality estimates for the East Africa data using methods based on double sampling of physically located patients who were lost to follow-up.[16 17] For South African data, we used similar methods, but lost patients were instead 'searched' in the death registries of the country.[28] Only passively determined death information was available in all other cohorts. For these cohorts, we used a modification of the approach by Brinkhof and colleagues,[5] which uses inflation factors to increase mortality under-reporting. Death rates based on observed data only were estimated for each patient subgroup defined by gender as well as age and CD4 category. These passive mortality estimates were then multiplied by a group-specific inflation factor determined as the ratio of the outreach-adjusted death rates over the observed rates estimated within the Kenyan and South African cohorts. This calculation resulted in two adjusted estimates of mortality—one using Kenyan data and the other using South African data—for cohorts where patient outreach was not available. This approach also addresses a number of limitations in the nomogram approach proposed by Egger and colleagues[38] in that it provides specific adjustments for all subgroups defined by gender, age and CD4 count as well as generating adjusted estimates past 1 year after ART initiation.

A final detail concerns the two-phase nature of the mortality hazard after initiation of ART. As has been widely published, the first 3–6 months after ART initiation are critical for the survival of the patient.[31 39] In this analysis, we used a piece-wise exponential model fit through Poisson regression. Four distinct intervals were defined at 0–6, 6–12, 12–24 and >24 months. This model assumes constant hazards of mortality in each interval (although these are allowed to vary from interval to interval). The period between 0 and 6 months was modelled separately from later periods to address the uniqueness of the early post-treatment period of high mortality.[39] Although this is a rather limited model in describing the changes in mortality which occurs after initiation of ART, it is useful because it is easy to implement and even easier to describe to stakeholders and decision makers involved with antiretroviral programmes around the world. Adjustments for dropout were handled by weighting according to the Frangakis and Rubin approach[19] to account for the increased HR between patients who were lost to follow-up versus those retained in care and under observation (see online supplementary appendix for further details of the modelling approach and the code used for model fitting).

RESULTS

This study encompasses five IeDEA regional cohorts from East, Southern and West Africa, Asia Pacific and Latin America. A total of 122 925 adult HIV-infected patients aged 15 years or older were included. From East Africa, the USAID-AMPATH cohort provided data from 32 217 adult patients while in the Southern Africa region, data from 48 046 patients were included from nine care and treatment cohorts in the Republic of South Africa. The West Africa IeDEA regional cohort was comprised of 15 cohorts from Benin, Côte d'Ivoire, Senegal, Gambia, Mali and Nigeria and contributed a total of 33 005

patients. The Asia Pacific regional cohort contributed 6106 patients from 13 countries including Cambodia, China, Hong Kong, Japan, India, Indonesia, Malaysia, the Philippines, Singapore, South Korea, Taiwan, Thailand and Vietnam. The regional cohort from Latin America contributed 3551 patients from Argentina, Brazil and Mexico.

The most commonly used first-line ART regimen in Kenya and South Africa is Stavudine (D4T) and Lamivudine (3TC) with either Nevirapine (NVP) or Efavirez (EFV). In Latin America, the most frequent combination is Zidovudine (AZT), 3TC and EFV. The second most common drug combination is Tenofovir (TDF) with Emtricitabine and EFV. In the Asia Pacific region, among developing countries, first-line therapy usually amounts to a combination of AZT or D4T plus 3TC with NVP or EFV, often as a generic drug similar to Africa, with a slow recent move towards TDF. In the developed countries (Japan, Hong Kong, Taiwan,

Table 1 Number of patients by gender and age category

	Gender	Age category				
Clinic site	**Females N (%)**	**15–24**	**25–34**	**35–44**	**45+**	**Total**
East Africa						
Kenya						
USAID-AMPATH	20 601 (63.9)	1765	11 359	11 942	7151	32 217
Southern Africa						
South Africa	27 791 (58.5)					
Aurum Institute (community)		661	4300	3953	2069	10 983
Aurum Institute (work place)		64	1231	1829	2164	5288
Gugulethu Clinic		233	1177	700	282	2392
Hlabisa HIV treatment and care programme		808	3424	2775	1742	8749
Khayelitsha ART Programme		667	3152	1841	653	6313
Masiphumelele Clinic		82	295	126	52	555
McCord Hospital		158	1185	918	427	2688
Themba Lethu Clinic		585	4199	3567	1644	9995
Tygerberg Hospital		90	483	343	167	1083
West Africa						
Benin						
CNHU Hubert H. Maga	496 (57.2)	56	310	309	192	867
Côte d'Ivoire	9738 (63.0)					
CePReF		243	1652	1520	826	4241
CNTS		65	455	417	220	1157
USAC, CHU de Treichville		151	1074	1093	644	2962
MTCT Plus		52	326	118	22	518
CIRBA		122	967	1127	767	2983
SMIT, CHU de Treichville		145	1377	2082	1140	4744
Burkina Faso						
CHU de Yaldago	1164 (68.8)	75	549	689	380	1693
Senegal	468 (55.8)					
SMIT Dakar		17	134	152	132	435
ANRS 1215, Dakar		23	136	155	90	404
Gambia						
Fajara	140 (64.5)	9	61	81	66	217
Mali	1403 (62.3)					
Gabriel Touré		146	583	516	318	1563
CHU de Point G		58	238	253	139	688
Nigeria	6756 (64.1)					
University of Abuja		435	2036	1519	1068	5058
University of Benin		298	1965	1523	1689	5475
Asia Pacific						
TAHOD	1232 (29.2)	290	1901	1375	658	4224
TASER	565 (30.0)	133	683	655	411	1882
Latin America						
Argentina						
Fundación Huésped	458 (30.3)	90	665	530	226	1511
Brazil						
Fundação Oswaldo Cruz	419 (33.3)	97	466	421	274	1,258
Mexico						
INNSZ	87 (11.1)	90	332	236	124	782

AMPATH, Academic Model Providing Access to Healthcare; ANRS, Agence nationale de recherché sur le sida; CHU, Centre Hospitalier Universitaire de Treichville (Côte d'Ivoire); CIRBA, Centre Intégré de recherché biocliniques d'Abidjan; CNHU, Centre Hospitalier Universitaire (Benin); CNTS, Centre national de transfusion sanguine; IeDEA, International Epidemiologic Databases to Evaluate AIDS; MTCT, mother-to-child transmission; SMIT, Service de Médecine Interne et Tropicale; TAHOD, TREAT Asia HIV Observational Database; TASER, TREAT ASIA Study to Evaluate Resistance; USAC, Unité de soins, d'accueil et de conseils pour les sidéens; USAID, US Agency for International Development.

Table 2 Loss to follow-up rates in the five regions involved in the analyses

	Male		Female		Total	
ART duration (months)	**N**	**Rate**	**N**	**Rate**	**Total**	**Rate**
East Africa						
0–6	1053	0.2340	1686	0.2079	2739	0.2172
6–12	394	0.1128	757	0.1189	1151	0.1168
12–24	434	0.0908	832	0.0951	1266	0.0941
>24	246	0.0765	466	0.0793	712	0.0783
South Africa						
0–6	1891	0.2331	2138	0.1796	4029	0.2013
6–12	665	0.1058	798	0.0852	1463	0.0935
12–24	913	0.1053	992	0.0782	1905	0.0892
>24	955	0.1230	815	0.0897	1770	0.1051
West Africa						
0–6	1799	0.3997	2687	0.3315	4486	0.3558
6–12	839	0.2314	1395	0.2108	2234	0.2181
12–24	1195	0.2206	2132	0.2149	3327	0.2169
>24	1603	0.2206	2590	0.2290	4193	0.2336
Asia Pacific						
0–6	176	0.0874	57	0.0679	233	0.0817
6–12	142	0.0785	55	0.0758	197	0.0676
12–24	238	0.0859	80	0.0628	318	0.0739
>24	654	0.0710	218	0.0604	872	0.0681
Latin America						
0–6	109	0.1104	51	0.1266	160	0.1151
6–12	96	0.1054	42	0.1113	138	0.1071
12–24	176	0.1087	79	0.1172	255	0.1112
>24	629	0.1314	252	0.1161	881	0.1266

ART, antiretroviral therapy.

Korea), first-line therapy would be a triple combination including an non-nucleoside reverse-transcriptase inhibitor or a protease inhibitor as in Europe, North America and Australia.

Baseline characteristics

A listing of the cohorts and the overall number of patients by gender and age category is given in table 1. The first impression from this table is that, while the epidemic in Africa is rather homogeneous and affects women in a roughly 2:1 ratio compared with men, the reverse is true in the Asia Pacific and Latin America where the epidemic is driven primarily by men. Loss-to-follow-up rates were high in all African cohorts (table 2). Annualized rates of loss to follow-up were over 20% during the first 6 months after ART initiation in the East and Southern Africa cohorts, with annualized rates of loss staying close to 10% during the second 6 months and stabilizing at that level thereafter. Losses to follow-up were almost double in West African cohorts, with annualized rates over 35% during the first 6 months after treatment initiation and remaining above 20% thereafter. Losses to follow-up were much lower in the Asia Pacific and Latin American cohorts, where annualized rates were 8% and just over 10% for patients after starting ART. These observations will be useful later when speculating about the applicability to these regions the results derived from cohorts in sub-Saharan Africa.

The distribution of CD4 count at ART initiation by gender and region is shown in figure 1. As it is widely reported,[7 40] women have higher CD4 counts at the time of ART initiation, the result of both early detection as well as complex healthcare access differences between the sexes, particularly in resource-limited settings.

Estimates of mortality

In the next subsection, we summarise the results from the two cohorts with available data on patient outreach (Kenya) and vital registry linkages (South Africa) as these data are amenable to mathematical adjustment of the mortality estimates. We proceed with a summary of the mortality estimates in the remaining three cohorts in the following subsection.

The Kenyan and South African cohorts

Mortality was estimated as described in the Methods Section at the critical first 6 months after the start of ART, 6–12, 12–24 and >24 months. Adjustments were based on 2409 successful outreach encounters in the Kenya cohort, out of 7252 patients who were lost to follow-up and 4602 individuals successfully linked out of 9167 persons who were lost to follow-up in the South African cohort. The adjusted estimates of mortality by age and CD4 category for male and female subjects are given in figure 2 for the Kenyan and South African cohorts. The immediate impression upon inspection of these figures is the high death rates in the first 6 months after initiation of ART and the significant reduction of these rates during subsequent periods. In order to appreciate the impact of the adjustments from information obtained through patient outreach (Kenya) and vital records (South Africa), we provide the correction factors defined as the ratios of the adjusted to the unadjusted death rates in each age, gender and CD4 category. These data are presented in figure 3. The implication of the results, as shown in figure 3, is that the adjustment of the death rates is both substantial and variable among genders, age groups, CD4 categories as well as cohorts. A major difference between the Kenyan and South African cohort is that, in the Kenyan cohort, the adjustments are the highest particularly during the first 6 months after ART initiation when they can reach 60%. Correction factors are even higher among South African cohorts where adjustments are almost 70% during the first 6-month period. In subsequent periods, the adjustments are even higher, sometimes over 100%, but their absolute effect is limited due to the low death rates during these periods. In the South African cohort, the highest adjustments involve the highest CD4 category (above 500 cells/μl). This suggests a possibly unique mortality pattern for this group.

The West African, Asia Pacific and Latin American cohorts

As mentioned in the Mathematical models section, many cohorts considered for this analysis have no outreach of patients and thus adjustments of patient mortality are impossible. We borrowed from the idea underlying the nomogram approach published by Egger and colleagues[31] but adjusted it to fit the needs of the current project. The nomogram approach considers the inverse relationship between the size of lost to follow-up (LTFU) and per cent of LTFU patients who are dead in order to extrapolate the correct (but unknown) death rate among patients who drop out from care. In the original paper by Egger and colleagues, estimates of 1-year mortality after initiating ART were produced.[31] However, here, we want to estimate mortality at 0–6, 7–12, 13–24 and >24 months from ART initiation in a large number of groups, where we cannot reasonably assume that the same correction factor holds. Thus, we used, as an estimate of the correction factor between LTFU patients and patients under observation in the paper by Egger and colleagues, the ratio of adjusted and crude death rates estimated in the East African (not shown) and South African cohorts for all the subgroups

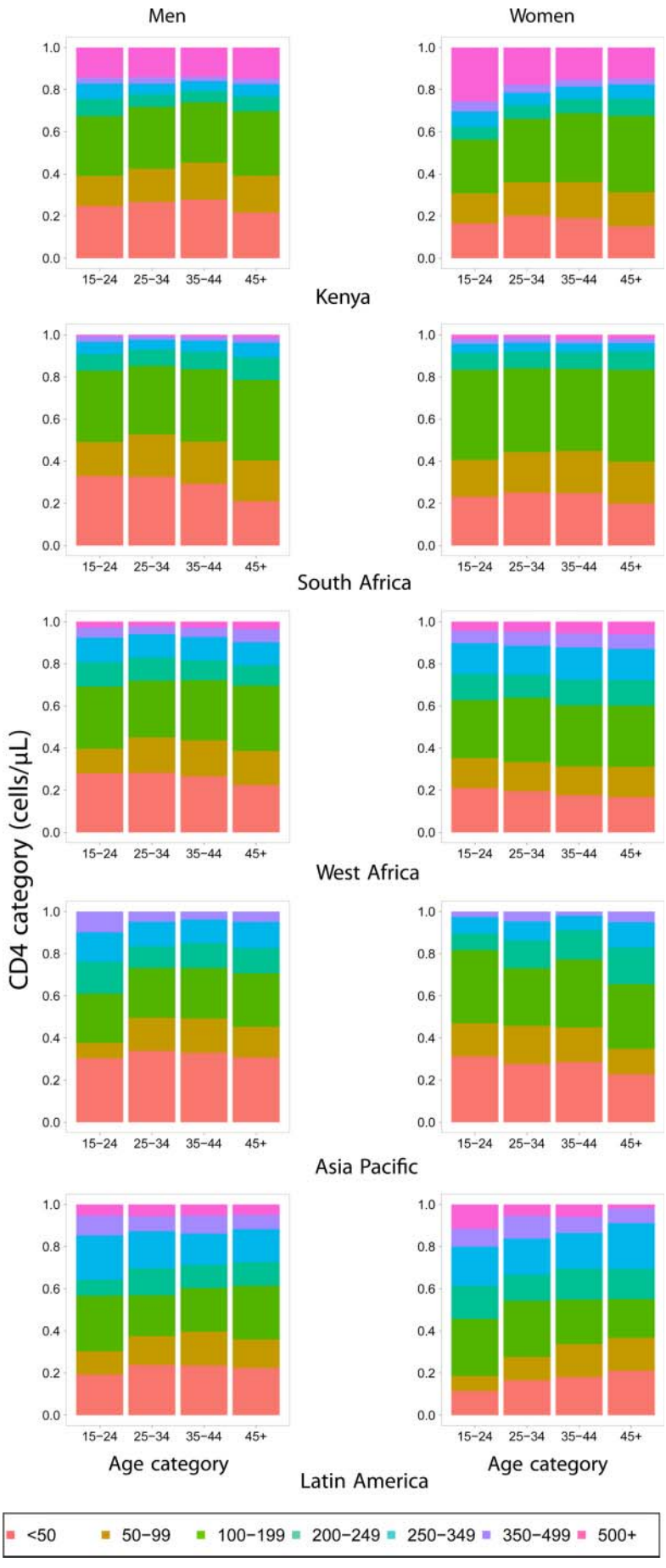

Figure 1 Distribution of CD4 counts at antiretroviral therapy initiation in the Kenya, South Africa, West Africa, Asia Pacific and Latin America cohorts. Statistics for men are shown on the left and women on the right.

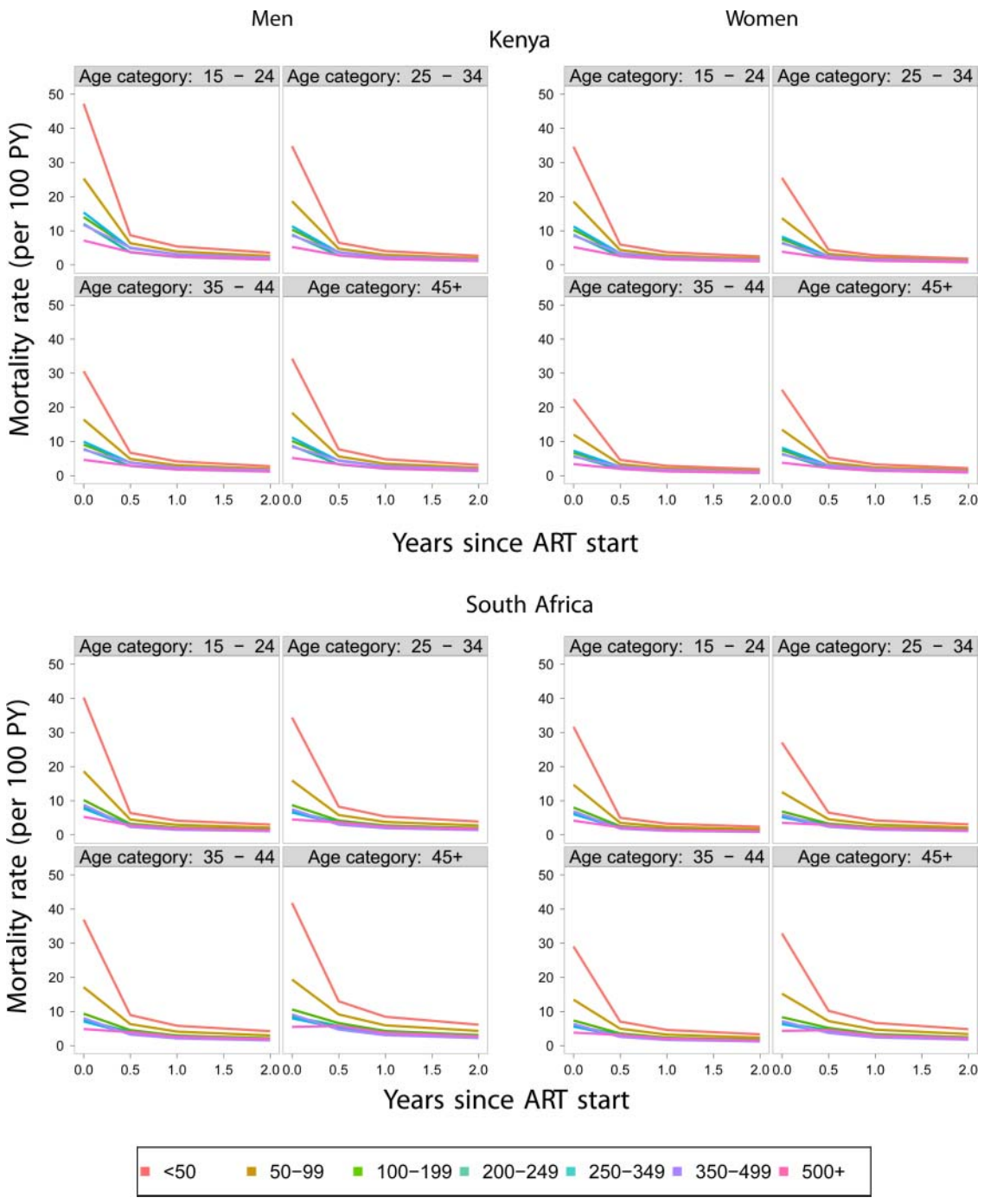

Figure 2 Adjusted estimates of mortality by age, gender and CD4 count antiretroviral therapy (ART) initiation in the Kenyan and South African cohorts.

(defined by gender, CD4 and age categories) considered in the current analyses (figure 3). The choice of the adjustment based on the South African data is justified further in the Discussion section. The revised estimates of mortality are shown in figure 4. In all cases, even though the highest death rates occur during the first 6 months after therapy initiation, estimated death rates are lower than in Kenya and South Africa. West Africa has the highest rates among these three regions, Latin America is in the middle and Asia Pacific has the lowest death rates among all regions considered.

DISCUSSION

This study summarises data from five regional consortia, three in sub-Saharan Africa, one in Latin America and one in the Asia Pacific region. As expected, the estimated death rates are the highest in sub-Saharan Africa, with moderately high rates in Latin America and low rates in Asia Pacific. As is widely published, use of outreach information and linkages with vital registries resulted in a substantial upward revision of all mortality estimates, particularly during the early months after initiation of ART.[16 17 19 24]

With respect to CD4 count at the start of therapy, the results are consistent with expectation. Patients who start therapy at lower CD4 counts, particularly <100 cells/µl, have significantly higher death rates compared with those starting at CD4 counts above 100 cells/µl. Men also have higher estimated mortality in all strata identified by gender, age and CD4 counts. A consistent trend emerges with regard to age as well.

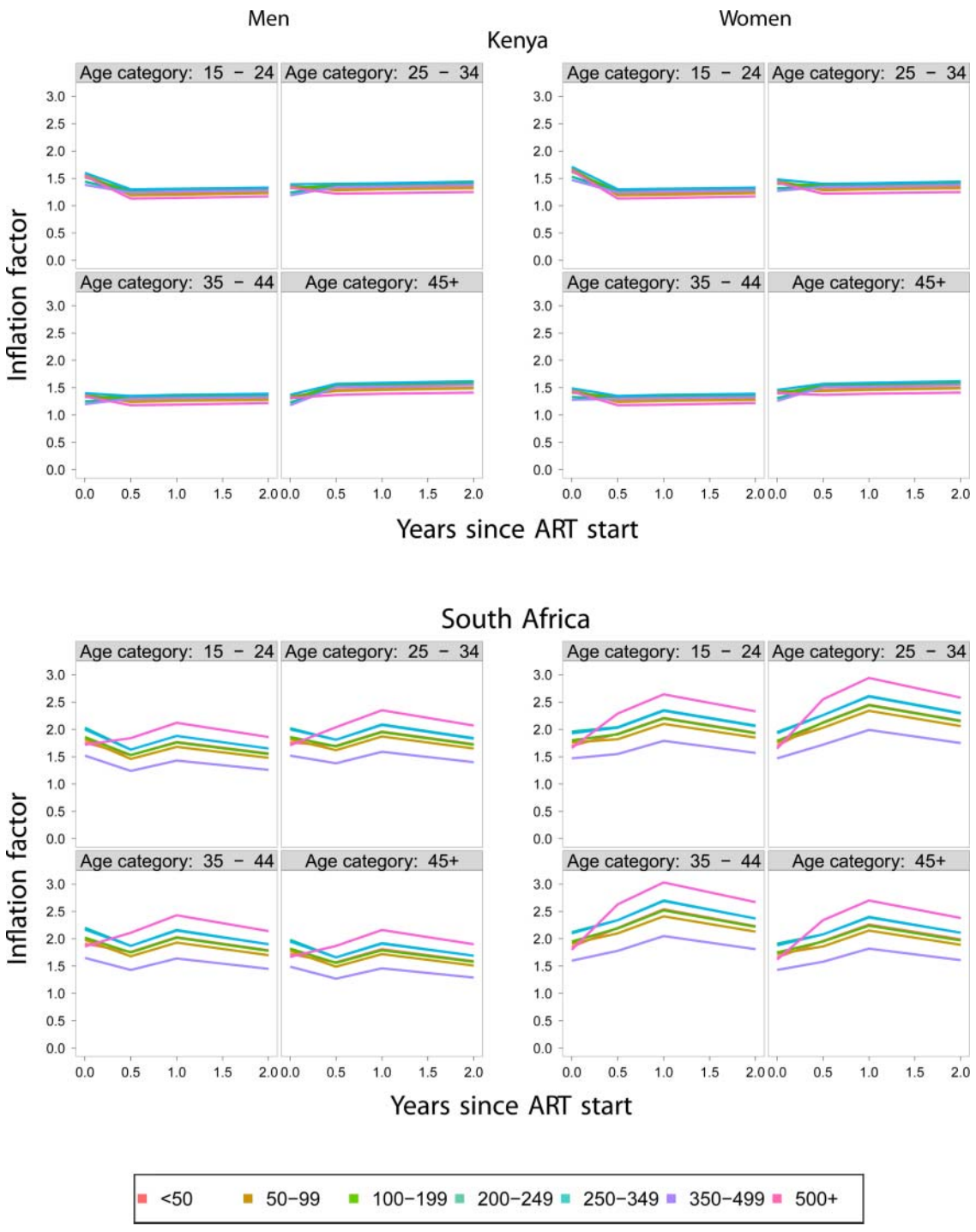

Figure 3 Correction factors for mortality estimates derived from the Kenya and South Africa cohort data. ART, antiretroviral therapy.

Mortality is highest among the 15–24 age group and the over-45 group and relatively lower in the other two groups.

Mortality adjustments did not change the overall trend of mortality. However, the magnitude of the revision was such as to significantly undermine the veracity of published estimates generated in a similar context but without the benefit of patient outreach or linkage to vital registries. This observation underlines the need of both the availability of post-dropout vital status information and the usefulness of mathematical adjustment methods to generate as accurate mortality estimates as possible. While mortality was, overall, lower in the highest CD4 group (>500 cells/μl), the adjustments in this category were the highest in the Southern Africa group. This suggests a separate mortality pattern in patients starting ART with higher CD4 counts, a not unsurprising result given that ART is initiated in these patients only when there is evidence of serious health problems. At the very least, this observation implies that current estimates of response to therapy (including mortality) cannot be used to predict what the impact of initiating ART at higher CD4 counts will be as advocated in newer treatment guidelines.

When outreach data are not available, our options are limited with regard to generating valid adjustments of mortality from passive recording of vital status information. Adjusting the

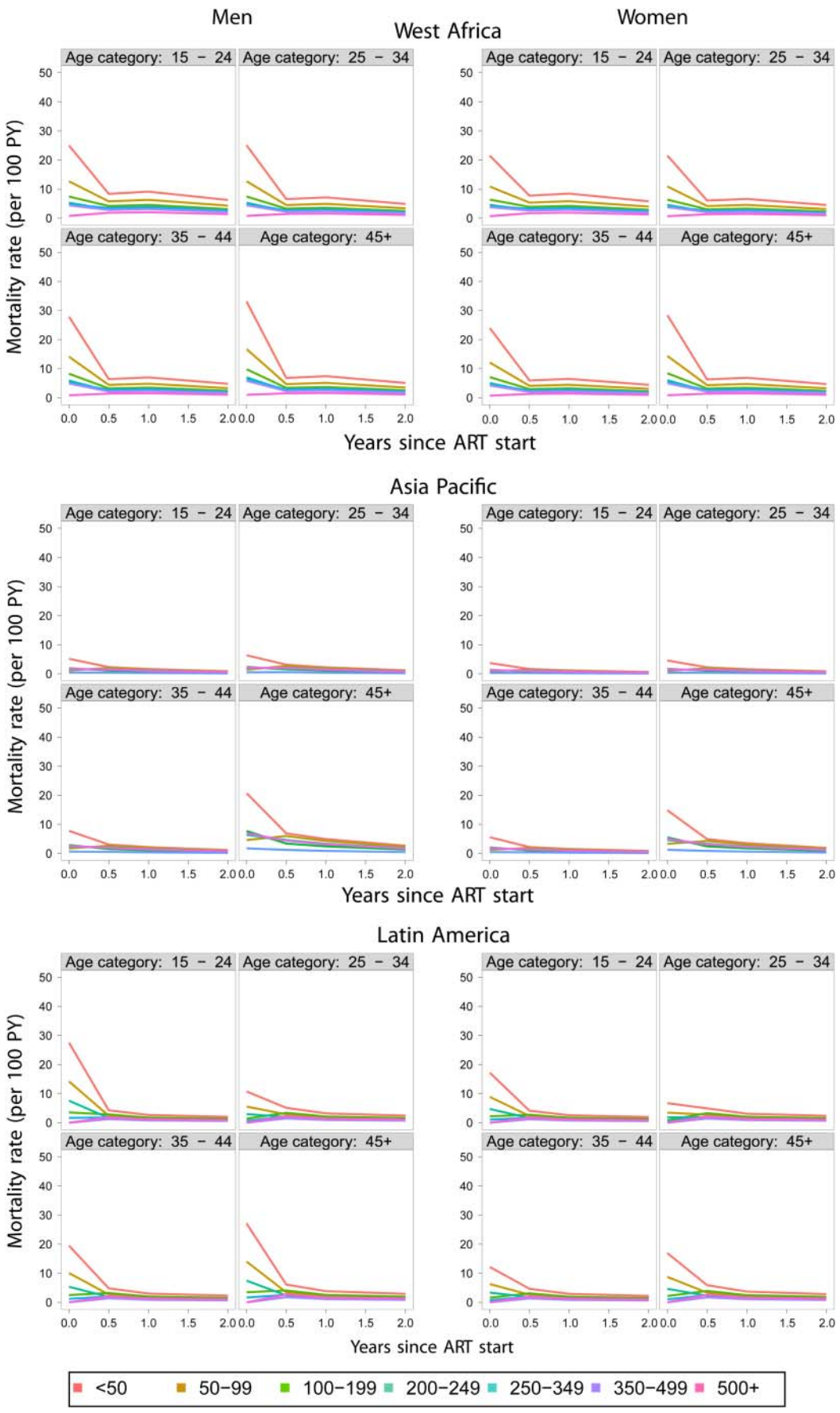

Figure 4 Adjusted death rates among men (left) and women (right panels) in West Africa, Asia Pacific and Latin America. ART, antiretroviral therapy.

analyses by incorporating predictive factors collected prior to dropout is not expected to work as the loss to follow-up is informative[41] (ie, it is related to the likelihood of the outcome even after adjusting for predictive factors that are commonly collected prior to dropping out). An alternative is to exploit the inverse relationship between the levels of dropout and the relative hazard between dropouts and non-drop outs[31] to impute the rate among individuals who are not retained in care. However, this approach requires significant contextual knowledge about the programme and may thus be questionable particularly in settings where high dropout is not associated with rapid scale-up and/or there are limited alternative care and treatment programmes in the area.[24]

To overcome this difficulty and derive mortality adjustments in the three regions without available patient outreach data (ie, the West Africa, Asia Pacific and Latin America regions), we used a hybrid method consistent with the approach in Brinkhof and colleagues[29] rather than the nomogram approach proposed by Egger *et al*.[31] Given the significant differences in the resulting adjustment when using the Kenya versus the South Africa correction factors, it is a valid question which of the two should be generally used. While, in absolute terms, the resulting mortality estimates were similar after employing either adjustment method, we ended up presenting the South Africa-derived correction in this manuscript. The reason for this is an idiosyncrasy of the Kenya cohort in which patients who have started ART more recently are followed particularly intensively compared with those who have been on treatment for longer periods.[16 17] This may have resulted in an undersampling in patients with a longer follow-up and may be the reason for the diminishing adjustment levels during the time intervals past 6 months. By contrast, access to vital registries is expected to be virtually unchanged regardless of the duration of therapy. This speculation along with the fact that both adjustment methods produced virtually identical adjustment coefficients during the first 6 months from the start of therapy resulted in us choosing the South Africa correction factors to adjust the mortality estimates in regions outside East and Southern Africa.

The approach applied here has a number of limitations: It is not known whether adjustments developed in sub-Saharan Africa will be applicable in settings outside the continent, particularly in developed countries such as Japan, Hong Kong or Taiwan. For example, patients who are lost to follow-up in Buenos Aires, Argentina (a middle-income setting), tended to have better markers of health at antiretroviral initiation than patients who remained in care,[42] which suggests that naive analyses could actually be overestimating death rates. In addition, rates of loss to follow-up are different between regions, something that the nomogram approach by Egger and colleagues[31] takes into account but our adjustment method does not. With that said, without available tracing data it is impossible to really know what is happening to those who are lost from observation, and the adjustment methods outlined here represent at least a reasonable starting point given the potentially significant under-reporting of death rates even in middle- or high-income settings.

However, even within regions with available outreach data biases exist. In East Africa, mortality adjustments were based on patient tracing programmes which do not trace a random sample of the patients who were lost to follow-up[19] but are imbedded into existing clinically-oriented patient outreach efforts.[18 19 25] The fact that the patients to be traced are not identified in a random fashion and, in addition, not all sought after patients are successfully located, raises significant questions about the representativeness of the located subsample compared with the overall group of patients lost from observations. The South African vital registry linkage methods are also subject to this bias. Thus, the accuracy of the adjusted estimates is put into question, particularly in light of reports where the hardest to locate patients are also those at the highest risk strata.[18] In the analyses presented in this paper we made the (largely untestable) assumption that the patients who are successfully located, either through physical outreach or successful linkage with the vital registry, do not differ systematically from those who are not and can thus be regarded as a random sample of all patients who are lost to follow-up.

A final consideration is the degree of confidence that one should attach to these results. We would expect that the adjustments produced in the Kenya and South African cohorts would be fairly accurate, particularly in the first 6 months from treatment initiation, given the enormity of the available data, which ensures that the mathematical properties of the models (ie, their 'asymptotics') would be fulfilled. Confidence in the mortality estimates would be decreasing in the case of West Africa and, even further, in the Latin America and Asia Pacific regions, which include highly variable epidemics. In particular, the better survival seen in the Asian data, which includes both developed and developing countries, should not be overinterpreted. The cohorts which comprise these data are not national treatment programmes, and survival in these relatively small numbers of patients in individual countries should probably not be extrapolated to survival outcomes across the entire region. For example, data from the China National Free Antiretroviral Treatment Program show survival more similar to outcomes in Africa.[43] The survival seen in the Asia data in this paper is perhaps better interpreted as indicative of what can be achieved in well resourced individual treatment sites.

The overall conclusion of this manuscript is thus twofold: From a clinical perspective, we continue to observe a large number of patients start therapy at very low CD4 counts, even in middle- and high-income settings, and these patients experience high levels of mortality particularly during the first 6-month period from treatment initiation. With this realisation, it is imperative to increase efforts to identify these patients and bring them to care. From an epidemiological and public-health perspective, patient outreach and vital linkage methods should be considered for inclusion in all HIV care and treatment programmes because of the strong association between loss to follow-up with adverse clinical outcomes and the resulting potentially significant underestimation of mortality.

Main findings

- Despite advances in antiretroviral treatment coverage, many patients start treatment at very low CD4 counts and experience significant mortality during the first 6 months after treatment initiation.
- Active patient tracing and linkage with vital registries are critical for deriving accurate estimates of mortality, particularly in low- and middle-income settings.
- In addition to the youngest patients, older patients have high death rates. With extended survival due to the expansion of ART coverage and the resulting ageing of the HIV/AIDS population, the impact of age on mortality needs to be considered in decision and policy making.

Acknowledgements We are grateful to all patients, care givers and data managers involved in the participating cohorts and treatment programmes. This work was supported by a contract provided by the Joint United Nations Programme on HIV/AIDS (UNAIDS). Individuals involved in the functioning of the programmes providing data to this study are provided in the Supplementary file online (http://dx.doi.org/10.1136/sextrans-2012-050658).

Contributors CTY, LFJ, AB, ME and TG conceived the study and performed all analyses. CTY wrote the manuscript. BSM conceived the database template. BSM, LFJ, BS, EB and ML created databases for analysis.

Funding This work was supported by UNAIDS through a grant to CTY, BSM, LFJ, AB, TG, EB and ME. In addition, the data collection was funded by the NIH National Institute of Allergies and infectious Diseases (NIAID), through grants AI069911 (East Africa), 2U01AI069924 (Southern Africa), AI069919 (West Africa), AI069907 (Asia Pacific) and AI069923 (Latin America).

Competing interests None.

Ethics approval Secondary analysis of anonymous primary clinical data.

Provenance and peer review Commissioned; externally peer reviewed.

REFERENCES

1. **UNAIDS.** *Global report: UNAIDS report on the global AIDS epidemic*. Geneva: Joint United Nations Programme on HIV/AIDS (UNAIDS), 2010.
2. **Braitstein P,** Brinkhof MW, Dabis F, *et al*. Mortality of HIV-1-infected patients in the first year of antiretroviral therapy: comparison between low-income and high-income countries. *Lancet* 2006;**367**:817–24.
3. **World Health Organization.** *Towards universal access. Scaling up priority HIV/AIDS interventions in the health sector. Progress report 2008*. Geneva: World Health Organization, 2008.
4. **Mills EJ,** Bakanda C, Birungi J, *et al*. Mortality by baseline CD4 cell count among HIV patients initiating antiretroviral therapy: evidence from a large cohort in Uganda. *AIDS* 2011;**25**:851–5.
5. **Brinkhof MW,** Boulle A, Weigel R, *et al*. Mortality of HIV-infected patients starting antiretroviral therapy in sub-Saharan Africa: comparison with HIV-unrelated mortality. *PLoS Med* 2009;**6**():e1000066.
6. **Brown ER,** Otieno P, Mbori-Ngacha DA, *et al*. Comparison of CD4 cell count, viral load, and other markers for the prediction of mortality among HIV-1-infected Kenyan pregnant women. *J Infect Dis* 2009;**199**:1292–300.
7. **Braitstein P,** Boulle A, Nash D, *et al*. Gender and the use of antiretroviral treatment in resource-constrained settings: findings from a multicenter collaboration. *J Womens Health (Larchmt)* 2008;**17**:47–55.
8. **Eckman A,** Huntley B, Bhuyan A. *How to integrate gender into HIV/AIDS programs: using lessons learned from USAID and partner organizations*. Washington, DC: Interagency Gender Working Group (IGWG) of USAID, 2004.
9. **Herstad B.** *Addressing gender issues related to HIV treatment adherence programs*. Washington, DC: Futures Group, USAID | Health Policy Initiative, Task Order 1, 2010.
10. **Rwechungura L,** Kayitare F. *Mapping of experiences of access to care, treatment, and support-Tanzania*. London: International Community of Women Living with HIV/AIDS (ICW), 2006.
11. **Orlando G,** Meraviglia P, Cordier L, *et al*. Antiretroviral treatment and age-related comorbidities in a cohort of older HIV-infected patients. *Hiv Med* 2006;**7**:549–57.
12. **Balestre E,** Eholie SP, Lokossue A, *et al*. Effect of age on immunological response in the first year of antiretroviral therapy in HIV-1-infected adults in West Africa. *AIDS* 2012;**26**:951–7.
13. **Rosen S,** Fox MP, Gill CJ. Patient retention in antiretroviral therapy programs in sub-Saharan Africa: a systematic review. *PLoS Med* 2007;**4**():e298.
14. **Shepherd BE,** Sterling TR, Moore RD, *et al*. Cross-cohort heterogeneity encountered while validating a model for HIV disease progression among antiretroviral initiators. *J Clin Epidemiol* 2009;**62**:729–37.
15. **Gerver SM,** Chadborn TR, Ibrahim F, *et al*. High rate of loss to clinical follow up among African HIV-infected patients attending a London clinic: a retrospective analysis of a clinical cohort. *J Int AIDS Soc* 2010;**13**:29.
16. **An MW,** Frangakis CE, Musick BS, *et al*. The need for double-sampling designs in survival studies: an application to monitor PEPFAR. *Biometrics* 2009;**65**:301–6.
17. **Yiannoutsos CT,** An MW, Frangakis CE, *et al*. Sampling-based approaches to improve estimation of mortality among patient dropouts: experience from a large PEPFAR-funded program in Western Kenya. *PLoS One* 2008;**3**():e3843.
18. **Anglaret X,** Toure S, Gourvellec G, *et al*. Impact of vital status investigation procedures on estimates of survival in cohorts of HIV-infected patients from Sub-Saharan Africa. *J Acquir Immune Defic Syndr* 2004;**35**:320–3.
19. **Geng EH,** Emenyonu N, Bwana MB, *et al*. Sampling-based approach to determining outcomes of patients lost to follow-up in antiretroviral therapy scale-up programs in Africa. *JAMA* 2008;**300**:506–7.
20. **Neyman J.** On the application of probability theory to agricultural experiments: essay on principles, section 9. *Transl Statist Sci 1990* 1923;**5**:465–80.
21. **Frangakis CE,** Rubin DB. Addressing an idiosyncrasy in estimating survival curves using double sampling in the presence of self-selected right censoring. *Biometrics* 2001;**57**:333–42.
22. **Robins JM,** Hernan MA, Brumback B. Marginal structural models and causal inference in epidemiology. *Epidemiology* 2000;**11**:550–60.
23. **Robins J,** Rotnitzky A, Bonetti M. Discussion of the Frangakis and Rubin article. *Biometrics* 2001;**57**:343–7.
24. **Henriques J,** Pujades-Rodriguez M, McGuire M, *et al*. Comparison of methods to correct survival estimates and survival regression analysis on a large HIV African cohort. *PLoS One* 2012;**7**:e31706.
25. **Tweya H,** Gareta D, Chagwera F, *et al*. Early active follow-up of patients on antiretroviral therapy (ART) who are lost to follow-up: the 'Back-to-Care' project in Lilongwe, Malawi. *Trop Med Int Health* 2010;**15**:82–9.
26. **Setel PW,** Macfarlane SB, Szreter S, *et al*. A scandal of invisibility: making everyone count by counting everyone. *Lancet* 2007;**370**:1569–77.
27. **Dorrington RE**, Moultrie TA, Timaeus IM. Estimation of mortality using the South African 2001 census data. Monograph 11. Centre for Actuarial Research, University of Cape Town. 2004. http://www.commerce.uct.ac.za/care/Monographs/Monographs/Mono11.pdf (accessed 29 Oct 2012).
28. **Van Cutsem G,** Ford N, Hildebrand K, Goemaere E, Mathee S, *et al*. Correcting for mortality among patients lost to follow up on antiretroviral therapy in South Africa: a cohort analysis. *PLoS ONE* 2011;**6**:e14684.
29. **Brinkhof MW,** Spycher BD, Yiannoutsos C, *et al*. Adjusting mortality for loss to follow-up: analysis of five ART programmes in sub-Saharan Africa. *PLoS One* 2010;**5**:e14149.
30. **Geng EH,** Bangsberg DR, Musinguzi N, *et al*. Understanding reasons for and outcomes of patients lost to follow-up in antiretroviral therapy programs in Africa through a sampling-based approach. *J Acquir Immune Defic Syndr* 2010;**53**:405–11.
31. **Egger M,** Spycher BD, Sidle J, *et al*. Correcting mortality for loss to follow-up: a nomogram applied to antiretroviral treatment programmes in sub-Saharan Africa. *PLoS Med* 2011;**8**:e1000390.
32. **Egger M,** Ekouevi DK, Williams C, *et al*. Cohort Profile: The international epidemiological databases to evaluate AIDS (IeDEA) in sub-Saharan Africa. *Int J Epidemiol* 2012;**41**:1256–64.
33. **McGowan CC,** Cahn P, Gotuzzo E, *et al*. Cohort Profile: Caribbean, Central and South America Network for HIV research (CCASAnet) collaboration within the International Epidemiologic Databases to Evaluate AIDS (IeDEA) programme. *Int J Epidemiol* 2007;**36**:969–76.
34. **Hamers RL,** Oyomopito R, Kityo C, *et al*. Cohort profile: The PharmAccess African (PASER-M) and the TREAT Asia (TASER-M) monitoring studies to evaluate resistance–HIV drug resistance in sub-Saharan Africa and the Asia-Pacific. *Int J Epidemiol* 2012;**41**:43–54.
35. **Gange SJ,** Kitahata MM, Saag MS, *et al*. Cohort profile: the North American AIDS Cohort Collaboration on Research and Design (NA-ACCORD). *Int J Epidemiol* 2007;**36**:294–301.
36. **Dabis F,** Balestre E, Braitstein P, *et al*. Cohort Profile: Antiretroviral Therapy in Lower Income Countries (ART-LINC): international collaboration of treatment cohorts. *Int J Epidemiol* 2005;**34**:979–86.
37. **Zhou J,** Kumarasamy N, Ditangco R, *et al*. The TREAT Asia HIV Observational Database: baseline and retrospective data. *J Acquir Immune Defic Syndr* 2005;**38**:174–9.
38. **Egger M,** Spycher BD, Sidle J, *et al*. Correcting mortality for loss to follow-up: a nomogram applied to antiretroviral treatment programmes in sub-Saharan Africa. *PLoS Med* 2011;**8**:e1000390.
39. **Yiannoutsos CT.** Modeling AIDS survival after initiation of antiretroviral treatment by Weibull models with changepoints. *J Int AIDS Soc* 2009;**12**:9.
40. **Kumarasamy N,** Venkatesh KK, Devaleenol B, *et al*. Regression of Kaposi's sarcoma lesions following highly active antiretroviral therapy in an HIV-infected patient. *Int J STD & AIDS* 2008;**19**:786–8.
41. **Rubin DB.** Inference and missing data. *Biometrika* 1976;**63**:581–92.
42. **Tuboi SH,** Schechter M, McGowan CC, *et al*. Mortality during the first year of potent antiretroviral therapy in HIV-1-infected patients in 7 sites throughout Latin America and the Caribbean. *J Acquir Immune Defic Syndr* 2009;**51**:615–23.
43. **Zhang F,** Dou Z, Ma Y, *et al*. Effect of earlier initiation of antiretroviral treatment and increased treatment coverage on HIV-related mortality in China: a national observational cohort study. *Lancet Infect Dis* 2011;**11**:516–24.

Acknowledgments:

We are grateful to all patients, care givers and data managers involved in the participating cohorts and treatment programmes. This work was supported by a contract provided by the Joint United Nations Programme on HIV/AIDS (UNAIDS). Individuals involved in the functioning of the programmes providing data to this study are as follows:

East Africa

USAID-AMPATH
Samwel Ayaya, Paula Braitstein, Vincent Ochieng-Ooko, Edwin Sang

Indiana University, Indianapolis, IN, USA
Kara Wools-Kaloustian, Beth Zwickl

South Africa

Aurum Institute for Health Research, Johannesburg, South Africa
Christopher Hoffmann

Hlabisa HIV Treatment and Care Programme, KwaZulu-Natal, South Africa
James Ndirangu

Khayelitsha ART programme, Cape Town South Africa
Daniela Belen Garone

Gugulethu and Masiphumelele Clinics, Cape Town, South Africa
Robin Wood

McCord Hospital, Durban, South Africa
Janet Giddy

Themba Lethu Clinic, Johannesburg, South Africa
Matthew Fox, Patrick MacPhail

Tygerberg Academic Hospital, Stellenbosch, South Africa
Hans Prozesky

University of Bern
Claire Graber, Fritz Kaeser

University of Cape Town
Mary-Ann Davies, Morna Cornell, Nicola Maxwell

West Africa

CNHU Hubert Maga, Cotonou, Benin

Djimon Marcel Zannou, Carin Ahouada, Jocelyn Akakpo, Christelle Ahomadegbé, Jules Bashi, Alice Gougounon-Houéto, Angèle Azon-Kouanou, Fabien Houngbé.

CHU de Yalgado, Ouagadougou , Burkina Faso:

Joseph Drabo, René Bognounou, Arnaud Dienderé, Eliezer Traore, Lassane Zoungrana, Béatrice Zerbo.

ACONDA-CePReF, Abidjan, Côte d'Ivoire

Eugène Messou, Joachim Charles Gnokoro, Mamadou Koné, Guillaume Martial Kouakou

MTCT-Plus, Abidjan, Côte d'Ivoire

Clarisse Amani Bosse, Kouakou Brou, Achi Isidore Assi

CIRBA, Abidjan, Côte d'Ivoire

Henri Chenal, Denise Hawerlander, Franck Soppi

CNTS, Abidjan, Côte d'Ivoire

Albert Minga, Yao Abo, Germain Bomisso

SMIT, CHU de Treichville, Abidjan, Côte d'Ivoire

Serge Paul Eholié, Mensah Deborah Noelly Amego, Viviane Andavi, Zelica Diallo, Frédéric Ello, Aristophane Koffi Tanon

USAC, CHU de Treichville, Abidjan, Côte d'Ivoire

Serge Olivier Koule, Koffi Charles Anzan, Calixte Guehi;.

CH Gabriel Toure, Bamako, Mali

Moussa Y Maïga, Fatoumata Fofana Diakité, Abdoulaye Kalle, Drissa Katile

CH Point G, Bamako, Mali

Hamar Alassane Traore*, Daouda Minta, Tidiani Cissé, Mamadou Dembelé, Mohammed Doumbia, Mahamadou Fomba, Assétou Soukho Kaya, Abdoulaye M Traoré, Hamady Traoré, Amadou Abathina Toure

UMB/IHV, Abuja, Nigeria

Man Charurat

UATH, Abuja, Nigeria

Samuel Ajayi, Georgina Alim, Stephen Dapiap, Otu

National Hospital Abuja, Nigeria

Festus Igbinoba

UBTH, Benin City, Nigeria
Okwara Benson, Clément Adebamowo, Jesse James, Obaseki, Philip Osakede

SMIT, CHU Fann, Senegal, Dakar:
Papa Salif Sow, Bernard Diop, Noël Magloire Manga, Judicael Malick Tine (SMIT, CHU Fann).

Fajara, Gambia
Kevin Peterson.

Asia Pacific (TAHOD- TASER)

University Malaya Medical Center, Kuala Lumpur, Malaysia
A Kamarulzaman, A Kajindran, LY Ong

Hospital Sungai Buloh, Sungai Buloh, Malaysia
C KC Lee, R David, and B LH Sim

National Center for HIV/AIDS, Dermatology and STDs, Phnom Penh, Cambodia;
CV Mean, V Saphonn, and K Vohith

Working Group on AIDS Faculty of Medicine, University of Indonesia/Ciptomangunkusumo Hospital, Jakarta, Indonesia
E Yunihastuti

Beijing Ditan Hospital, Capital Medical University, Beijing, China
FJ Zhang, HX Zhao, N Han

Division of Infectious Diseases, Department of Internal Medicine, Yonsei University College of Medicine, Seoul, South Korea
JY Choi, SH Han, JM Kim

Hospital Raja Perempuan Zainab II, Kota Bharu, Malaysia
M Mustafa, N Nordin

YRG Centre for AIDS Research and Education, Chennai, India
N Kumarasamy, S Saghayam, C Ezhilarasi

Tan Tock Seng Hospital, Singapore;
OT Ng, A Chua, LS Lee, A Loh

Queen Elizabeth Hospital, Integrated Treatment Centre, Hong Kong, China
PCK Li, MP Lee, KH Wong

Chiang Rai Prachanukroh Hospital, Chiang Rai, Thailand;
P Kantipong, P Kambua

HIV-NAT/Thai Red Cross AIDS Research Centre, Bangkok, Thailand
P Phanuphak, K Ruxrungtham, M Khongphattanayothin, S Sirivichayakul

Research Institute for Tropical Medicine, Manila, Philippines
R Ditangco, E Uy, R Bantique

National Center for Global Health and Medicine, Tokyo, Japan
S Oka, J Tanuma, T Nishijima

Institute of Infectious Diseases, Pune, India
S Pujari, K Joshi, A Makane

Faculty of Medicine, Ramathibodi Hospital, Mahidol University, Bangkok, Thailand
S Sungkanuparph, S Kiertiburanakul, L Chumla, N Sanmeema

Faculty of Medicine, Udayana University and Sanglah Hospital, Bali, Indonesia
TP Merati†, DN Wirawan, F Yuliana

Research Institute for Health Sciences, Chiang Mai University, Chiang Mai, Thailand
T Sirisanthana, R Chaiwarith, W Kotarathititum, J Praparattanapan

Bach Mai Hospital, Hanoi, Vietnam
TT Pham, DD Cuong, HL Ha

National Hospital for Tropical Diseases, Hanoi, Vietnam
VK Nguyen, VH Bui, TT Cao

Faculty of Medicine, Siriraj Hospital, Mahidol University, Bangkok, Thailand
W Ratanasuwan, R Sriondee

Taipei Veterans General Hospital and AIDS Prevention and Research Centre, National Yang-Ming University, Taipei, Taiwan;
YMA Chen, WW Wong, YJ Chen, LH Kuo, YT Lin

Brown University, Providence, RI, USA
R Kantor

TREAT Asia, amfAR - The Foundation for AIDS Research, Bangkok, Thailand
AH Sohn, N Durier, B Petersen, T Singtoroj

The Kirby Institute, University of New South Wales, Sydney, Australia.
DA Cooper, A Jiamsakul,

Latin America

Fundación Huésped, Buenos Aires, Argentina
Carina Cesar and Pedro Cahn

Fundação Oswaldo Cruz, Rio de Janeiro, Brazil
Ronaldo Ismerio, Valdilea Gonçalves Veloso, Beatriz Grinsztejn and Paula Mendes Luz

Instituto Nacional de Ciencias Médicas y Nutrición, Salvador Zubirán, Mexico City, Mexico
Yanink Caro-Vega, Brenda Crabtree-Ramirez, and Juan Sierra-Madero from Mexico

Vanderbilt University, Nashville, TN, USA
Catherine McGowan

STATISTICAL APPENDIX

MODELS

We present here some mathematical details of the model used in the adjusted analyses. The general exponential regression model of survival is a proportional hazards model of the hazard (the instantaneous risk of death). The general proportional hazards model is as follows:

$$\lambda_i(t) = \lambda_0(t)\exp(\boldsymbol{x}_i'\boldsymbol{\beta})$$

where the vector $\boldsymbol{x}$ contains the measurements of interest (in our case, gender, CD4 count and age category at ART initiation), and $\lambda_0(t)$ is the "baseline" hazard (the hazard of a person with all risk factors set at their reference values). The exponential survival model has the added simplification that $\lambda_0(t) = \lambda_0$ for all time points t (constant hazards) so that the exponential model for the hazard is

$$\lambda_i(t) = \lambda_0\exp(\boldsymbol{x}_i'\boldsymbol{\beta})$$

Since the constant-hazard assumption is rather restrictive, we fit a *piecewise* exponential model where $\lambda_0(t) = \lambda_{0j}$ for t in the time interval $[\tau_{j-1}, \tau_j)$ with $\tau_0 = 0, \tau_1 = 6, \tau_2 = 12, \tau_3 = 24, \tau_4 = \infty$ and $j = 1,2,3,4$. Put more simply, we assumed that the baseline hazard of death is constant in the four time intervals (in months from ART start) $[0,6), [6,12), [12,24)$ and ≥ 24 months.

As mentioned in the methods section in the body of this document, we fit two separate models, one for the time $[0,6)$ (ART initiation to six months) and one for the three subsequent time intervals (i.e., after 6 months). This was because neither the hazard during the first six months after the start of therapy, nor the associations of the predictive factors (gender, CD4 count and age) are expected to follow the same pattern with subsequent periods (see Yiannoutsos et al., for a more mathematical treatise of change points in hazards during the period after initiation of therapy[1]). We will come back to this later on. For now, we keep considering the simpler (single) Exponential model as stated above.

Clearly, an adjustment needs to be made to account for the fact that a large proportion of the patient cohort, a subgroup including some patients with very adverse prognosis, has been lost from observation. This is because we cannot consider the subgroup of patients who remain on observation as representative of the dropouts, since patients who are lost have been observed to have much higher mortality hazard[2-4]. To do this we consider the vital status ascertained on a subset of the lost patients (which we consider a random sample of all patient dropouts) and use this information to update the hazard for all dropouts. In the case of the AMPATH cohort, the vital status of this subset was located through tracing of patients who missed scheduled visits (see Yiannoutsos et al., for a description of this program[4]). For South African cohorts, the vital status of lost patients was established through linkages

to the national death registry[5]. In both cases, the subset of patients with sufficient civilian information to be successfully located was assumed to be representative (i.e., to form a random sample) of the entire lost to follow-up cohort.

Following the approach by Frangakis and Rubin[6] and An et al.[2], we generate an average of the mortality hazard weighted by the inverse proportion of the number of patients who drop out (numerator) versus the number who were located (the random sample; denominator). This has the effect of multiplying each of the located patients by a number which is larger than one, thus creating virtual "copies" of these individuals to replace those whose vital status was not ascertained (as the traced patients are considered to be representative for the individuals who are lost and not traced). We replace the remainder of the lost patients who were not located by these copies, while the data from the actual individuals (who dropped out but did not have their vital status ascertained) are weighted by zero (effectively being excluded from the analysis). Patients who remained on observation are weighted by one (i.e., they only represent themselves). Their weighted hazard is thus

$$\lambda(t) = \sum_{g=0}^{1} w_g \lambda_g(t)$$

with g representing dropouts $(g = 1)$ and non-dropouts ($g = 0$) and

$$w_g = \begin{cases} \frac{n_0}{\tilde{n}_0} & \text{if dropout and traced} \\ 1 & \text{if non} - \text{dropout} \end{cases}$$

where n_0 is the total number of dropouts and $\tilde{n}_0$ is the subset of dropouts who were traced.

Transferring the same idea to our piecewise exponential regression model, we have

$$\lambda_i(t) = \sum_{g=0}^{1} w_g \lambda_{0g}(t) \exp(\boldsymbol{x}_i'\boldsymbol{\beta}) = \left\{ \sum_{g=0}^{1} \sum_{j=1}^{4} w_g \lambda_{0jg} I\left[t \in [\tau_{j-1}, \tau_j)\right] \right\} \exp(\boldsymbol{x}_i'\boldsymbol{\beta})$$

for $j = 1,2,3,4$ and $I[\cdot]$ is an indicator function.

Now, returning to the fact that a separate model was fit for the time interval $[\tau_{j-1}, \tau_j) = [0,6)$ and the last three time intervals $[\tau_{j-1}, \tau_j)$ for $j = 2,3,4$ i.e., $[6,12)$, $[12, 24)$ and $[24, \infty)$, we obtain the final model used for the analyses presented in this paper

$$\lambda_i(t) = \sum_{g=0}^{1} \left\{ w_g \lambda_{01g} I[t \in [\tau_0, \tau_1)] \exp(\boldsymbol{x}_i'\boldsymbol{\beta}_1) + \sum_{j=2}^{4} w_g \lambda_{0jg} I\left[t \in [\tau_{j-1}, \tau_j)\right] \exp(\boldsymbol{x}_i'\boldsymbol{\beta}_2) \right\}$$

Note that two different regression coefficients $\boldsymbol{\beta_1}$ and $\boldsymbol{\beta_2}$ are calculated in these two models, corresponding to different association between the risk factors before and after six months from ART initiation respectively.

The two-part piece-wise exponential model described above was fit by the equivalent Poisson log-linear model on the constant hazards[7 8]. The analysis was implemented by the STATA software version 11.1 (StataCorp, College Station, TX, USA). The program code is given in the following Appendix.

STATA ANALYSIS

Description of the database

The following code produced the analysis presented in this white paper. It assumes a data set where the following minimum number of variables are present:

Variable name	Type	Format	Description
ptidno	long	8.0	Patient ID number
dob	date		Date of birth
dod	date		Date of death
lastvisitdt	date		Date of last visit
death	byte	8.0	Death indicator
ageatarvstart	float	8.0	Age at the start of ART
cd4	Integer	8.0	CD4 count (cells/μL)
ltfu	byte	8.0	Lost to follow-up indicator
arvstartdt	date		Date of ART initiation
oraftervis	byte	8.0	Patient outreach indicator

Here are some comments before we proceed to the code:

1. `oraftervis` is equal to 1 ("yes") if the dropout patient was located by outreach. If not, then `oraftervis`=0 and we assume that no attempt was made (i.e., we equate patients who were not outreached with those who were outreached but not found). This is an important limitation of this study. A similar set of assumptions is made with respect to patients without sufficient information for linkage with a vital registry (they are considered as not having been found).
2. The last visit date `lastvisitdt` includes, for patients who were located via outreach, the last date known to be alive (so the last "visit" for these patients is their last contact with site outreach staff). In linkages with the vital registry, this date is the most recent date of inquiry.
3. Different lost-to-follow-up definitions produce slightly different results, as different numbers of patients are declared lost. This is not expected to be of major concern.
4. This study also assumed that all death dates (`dod`) were exact. However, some death dates were estimated (mainly due to vital status information obtained by proxy). It is tacitly assumed

that these estimates do not include a systematic error (i.e., they have random variations around the unknown true death date)

5. The indicator `death` includes all deaths, both those ascertained through passive/routine means and those ascertained through active means (i.e., patient outreach).

STATA analysis code

```
* Ensure that last visit date is updated for date of death

gen new_lastvisitdt=max(dod,lastvisitdt) if dod ~=.
replace new_lastvisitdt=lastvisitdt if dod==. & death==0

preserve

* split for duration: origin and entry are first ART visit
stset new_lastvisitdt , fail(death==1) scale(365.25) id(ptidno) /*
*/ enter(arvstartdt) origin(arvstartdt)

stsplit durcat , at(0.5 1 2)

* split for age: origin is dob, entry is arvstartdt

stset new_lastvisitdt , fail(death==1) scale(365.25) id(ptidno) /*
*/      enter(arvstartdt) origin(dob)

stsplit new_agecat , at(15 25 35 45)

* * * * * * * * * * * * * * Weighted analysis * * * * * * * * * * * * *
* Create the weights

gen weight=1

* Count deaths among LTFU or outreached patients
count if (oraftervis ==0 & ltfu ==1 & death~=1)| (oraftervis ==1)

* Create weights for outreached (i.e., located) patients only

replace weight =r(N) if (oraftervis ==1)
count if oraftervis ==1
replace weight =weight/r(N) if oraftervis ==1

* Exclude dropouts who were not outreached from the analysis

drop if oraftervis ==0 & ltfu ==1 & death~=1


* The following code is a check that weights fill in excluded patients

sum weight
di r(mean)*r(N)

* The result of the above calculation must equal to the original N!

* reset to have origin at arvstart with pweight=weight

stset new_lastvisitdt [pweight=weight], id(ptidno) failure(death==1) /*
*/ enter(time arvstart)  origin(time arvstart) scale(365.25)

* reset to have origin at arvstartdt
```

```
stset new_lastvisitdt , fail(death==1) scale(365.25) id(ptidno) /*
*/ enter(arvstartdt) origin(arvstartdt)

* Change time scale per 100 person-years

gen pyo100=(_t-_t0)/100
```

```
*************************** Major point ***************************
* Note that, when intervals get split, STATA does not update the
* death indicator correctly.  The death indicator must be zero in
* all but the final interval and equal to the original death indicator
* in the last interval.
* Thus, we need to use the STATA internal event indicator _d in the
* calculations instead of the original death indicator because
* (correctly) _d=0 for all intervals prior to the last one and _d=death
* at the last interval.
* Thus, an analysis involving the death variable would be wrong!
*********************************************************************

* Poisson models

* Model for the first 6 months since ART initiation
xi: poisson _d i.male i.new_agecat i.cd4cat if(pyo100>0 & durcat==0), /*
*/  exposure(pyo100)

* Model for after the first six months since ART initiation
xi: poisson _d i.male i.new_agecat i.cd4cat i.durcat  /*
*/  if(pyo100>0 & durcat>0), exposure(pyo100)

* * * * * * * * * * End of weighted analyses * * * * * * * * * * *
```

References

1. Yiannoutsos CT. Modeling AIDS survival after initiation of antiretroviral treatment by Weibull models with changepoints. *J Int AIDS Soc* 2009;12(1):9.
2. An MW, Frangakis CE, Musick BS, Yiannoutsos CT. The need for double-sampling designs in survival studies: an application to monitor PEPFAR. *Biometrics* 2009;65(1):301-6.
3. Egger M, Spycher BD, Sidle J, Weigel R, Geng EH, Fox MP, et al. Correcting mortality for loss to follow-up: a nomogram applied to antiretroviral treatment programmes in sub-Saharan Africa. *PLoS Med* 2011;8(1):e1000390.
4. Yiannoutsos CT, An MW, Frangakis CE, Musick BS, Braitstein P, Wools-Kaloustian K, et al. Sampling-based approaches to improve estimation of mortality among patient dropouts: experience from a large PEPFAR-funded program in Western Kenya. *PLoS One* 2008;3(12):e3843.
5. Van Cutsem G, Ford N, Hildebrand K, Goemaere E, Mathee S, Abrahams M, et al. Correcting for Mortality Among Patients Lost to Follow Up on Antiretroviral Therapy in South Africa: A Cohort Analysis. *PLoS One* 2011;6(2).
6. Frangakis CE, Rubin DB. Addressing an idiosyncrasy in estimating survival curves using double sampling in the presence of self-selected right censoring. *Biometrics* 2001;57(2):333-42.
7. Holford TR. The analysis of rates and of survivorship using log-linear models. *Biometrics* 1980;36(2):299-305.
8. Laird N, Olivier D. Covariance analysis of censored survival data using log-linear analysis techniques. *J Am Stat Assoc* 1981;76:231-40.